PCT / CA 00/00738
09 AUGUST 2000 (09.08.00)

CA00/00738

REC'D 05 SEP 2000
WIPO PCT

PA 269435

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

4

July 05, 2000

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: ***60/139,788***
FILING DATE: ***June 21, 1999***

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

H. PHILLIPS
Certifying Officer

PROVISIONAL APPLICATION COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(b)(2)

Docket Number	11004-1	Type a Plus Sign (+) Inside this Box ->	+

INVENTOR(S)/APPLICANT(S)

LAST NAME	FIRST NAME	Middle Initial	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
JAIN	Ravinder	K.	2413 Irvine Avenue, Saskatoon, Sask. S7J 2A9, Canada
MACKENZIE	Samuel	L.	17 Cambridge Crescent, Saskatoon, Sask. S7H 3P9, Canada

TITLE OF THE INVENTION (280 CHARACTERS MAX.)

Oil Content And Composition enhancement by Targeted/Non-Targeted Expression Of Plant Plastidial And Bacterial Gene For Glycerol-3-Phosphate Acyltransferase

CORRESPONDENCE ADDRESS

J. Wayne Anderson
National Research Council of Canada
Intellectual Property Services Office, EG-10, Bldg. M-58
Montreal Road, Ottawa, Ontario, Canada K1A 0R6

STATE	Ontario	ZIP CODE	K1A 0R6	COUNTRY	Canada

ENCLOSED APPLICATION PARTS (*Check all that Apply*)

X	*Specification*	*Number of pages*	10			*Small Entity Statement*
X	*Drawing(s)*	*Number of Sheets*	8			*Other (specify)*

METHOD OF PAYMENT (*Check One*)

	A check or money order is enclosed to cover the Provisional filing fees	*Provisional Filing Fee Amount ($)*	$ 150.00
X	*The Commissioner is hereby authorized to charge filing fees and credit Deposit Account Number* 14-0429		

The invention was made by an agency of the United States Government or under contract with an agency of the United States Government. X No

Date: 21 June 1999.

Respectfully submitted,

J. Wayne Anderson
Patent Agent for Applicant
Regn No. 28,158

Enclosurers
:ech

_____ Additional inventors are being named a separately numbered sheets attached hereto.

PROVISIONAL APPLICATION FILING ONLY

OIL CONTENT AND COMPOSITION ENHANCEMENT BY TARGETED/NON-TARGETED EXPRESSION OF PLANT PLASTIDIAL AND BACTERIAL GENE FOR GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE

There is a general industrial demand for higher oil content in living organisms in order to be more cost effective during production and processing. It is estimated that a 1% increase in seed oil would add $20 M to the value of the Canadian canola crop at the farm gate.

One way of accomplishing it is by modifying metabolic pathways of living organisms by enhancing the expression of genes involved in oil synthesis. For example, overexpression of a yeast SLC1-1 has been found to increase oil content in plants (Zou et al.,[1] 1997).

Only a limited number of genes have been reported to enhance oil content (for example, Zou et al.[1], 1997, Marillia et al.[2], 1998). There is a need to explore other approaches using other genes which might be more effective alone or in combination with previously cloned genes.

We propose to increase oil content and modify oil composition by overexpressing the genes for the enzyme glycerol-3-phosphate acyltransferase (GPAT). This has been achieved in yeast. Details of the procedures and findings are provided below.

This is the first demonstration of an increase in carbon flux into triglycerides by manipulating the expression of a GPAT enzyme.

The invention is immediately applicable to the improvement of processes in which commercially important fatty acids (e.g. gamma linolenic acid) and cocoa butter are produced in yeast.

BACKGROUND

GPAT is the first enzyme of membrane and storage lipid (triacylglycerol; TAG) biosynthesis in living organisms. It determines the type of fatty acid incorporated in the *sn*-1 position of membrane and storage lipids. Overall, its specificity for various acyl-CoA appears to be flexible (Frentzen, [3] 1993; Frentzen and Wolter, [4] 1998). Since it is the first step in triglyceride and phospholipid lipid biosynthesis, it is surmised that enhancing production of GPAT might also result in an increase of 1-acylglycerol-3-phosphate through the Kennedy pathway, resulting in higher oil content. Three types of GPAT have been reported from plants: plastidial (P), mitochondrial (M), and cytosolic (ER). They exhibit variable specificity toward acyl-CoA and acyl-ACP substrates (Frentzen, [3] 1993). The ER form of GPAT is considered to be the most important for TAG manipulation but it has not been possible to isolate a gene for this enzyme. However, cDNAs for P-GPATs have been isolated and characterized from a number of species. Also, heterologous GPAT genes from bacteria and mammals are available (see Wilkison and Bell [5], 1997; Dircks and Sul, [6] 1997).

A plastidial GPAT gene has been used by others to change the fatty acid composition of membrane lipids and to improve chilling tolerance (Murata, et al, [7] 1992 and Nishizawa, [8] 1996). The bacterial GPAT gene (*plsB*) has also been used to change the fatty acid composition of membrane lipids and to decrease chilling tolerance (Wolter et al, [9] 1992). None of the studies thus far have used GPAT genes to increase oil content in a living organism.

Because of the unavailability of an ER GPAT gene from plants, we set out to use a plant plastidial GPAT and a bacterial GPAT gene to establish the proof of our concept. Here we demonstrate the use of a plant plastidial or bacterial GPAT gene to increase oil content. Two genes were modified to create four versions to achieve higher oil content.

A cDNA (~gene) for a P-GPAT has been isolated in this lab from safflower (*Carthamus tinctorius*) (Bhella and MacKenzie, [10] 1994). A P-GPAT can use both acyl-ACP and acyl-CoA substrates, although the latter with lesser efficiency (Frentzen, [9] 1993). It was surmised that in the absence of any acyl-ACP in the cytosol, a P-GPAT will have to use acyl-CoA, although it has to compete with the ER-GPAT for substrates, acyl-CoAs and glycerol-3-phosphate (G-3-P). Its competitive ability could be enhanced by overexpressing it by using a strong constitutive or a tissue specific promoter. Normally, a P-GPAT is targeted to plastids by a transit peptide (TP). The TP was deleted to confine the GPAT activity in the cytosol (P-GPAT-TP). The modified gene was labelled as *ctpgpat-tp*. It is anticipated that an ER environment may also be required for a P-GPAT-TP to effectively participate in TAG biosynthesis. The ER retention signals (ERRS; Jackson et al., [11] 1990) were added to P-GPAT-TP to meet this requirement. The modified gene was labelled as *ctpgpat-tp+errs*.

The *Escherichia coli's plsB* gene has been previously characterized (see Wilkison and Bell, [5] 1997). The plsB can use both acyl-ACP and acyl-CoA equally effectively. The *plsB* gene was used as such and also with an added ERRS sequence. The genes were labelled as *plsB* and *plsB+errs*.

EXPERIMENTAL

An open reading frame (*orf*) without *tp* (~1.1kb) was PCR amplified from the *ctpgpat* cDNA. Another chimeric gene containing an *errs* at its 3' end was also PCR amplified. The *orf* of the *E. coli plsB* gene (~2.5kb) was also PCR amplified as it is or with an *errs* at its 3' end. The blunt-end PCR fragments generated using *Pfu* DNA polymerase were cloned into pSK II cloning vector and were sequenced to confirm the incorporation of restriction sites and *errs* sequences to chimeric genes. The *ctpgpat* or *plsB* chimeric genes were retrieved as *Bam*HI or *Bgl*II segments, respectively, and were cloned into *Bam*HI site of the yeast expression vector, pYES2 (*Invitrogen*), under the

control of a galactose inducible promoter. The INVSc strain (*Invitrogen*) of yeast was transformed with different recombinant constructs.

Yeast cells containing chimeric GPAT genes were grown in a medium containing glucose. GPAT gene expression was induced by transferring cells to growth medium containing galactose. Total proteins form the control and GPAT overexpressing cells were extracted and production of GPAT protein by genes was tested by doing GPAT enzyme assay (Eccleston and Harwood, [12] 1995). The products of the reaction were separated into lyso-phosphadic acid (LPA), phosphatidic acid (PA), di-acyl glycerol (DAG) and TAG by thin layer chromatography. Total lipids in control and cells in which GPAT was overexpressed were measured by ^{1}H-NMR (Rutar, [13] 1989) and by gas liquid chromatography.

RESULTS AND DISCUSSION

The modified genes encoded for GPAT activity as shown by an *in vitro* enzyme assay (Table 1). Also, increased GPAT activity led to more lipid production in an *in vitro* assay (Table 2). However, *in vitro* results may not necessarily reflect *in vivo* synthesis as extra substrate is provided in the *in vitro* assay. Nevertheless, we discovered that more lipids were also produced *in vivo* by enhancing the expression of GPAT genes (Tables 3&4).

Here, the proof of the concept was shown in yeast, a eukaryotic model system. Since GPAT is a ubiquitous enzyme for lipid biosynthesis in all living organisms, it is reasonable to expect that the overexpression of these genes will also enhance oil synthesis in other microorganisms and plants. Moreover, in *in vitro* assays, the safflower plastidial GPAT prefers unsaturated fatty acids as compared to *E. coli* GPAT which prefers saturate fatty acids (Table 1). These genes can therefore be used to modify the type of fatty acid at the *sn*-1 position of triglycerides. This will lead to production of structured triglycerides in which the fatty acids occupying each position are limited to only one, or at most, a few specific fatty acids. There is literature evidence

that fatty acid absorption and physiological effect are related to triglyceride structure, not just gross composition.

Alternatively, the inhibition of GPAT activity by antisense may reduce fats in mammals as higher GPAT activity has been found to be associated with obese rats than with lean rats (Jamadar et al., [14] 1996). Inhibition of GPAT could also be used in plants to reduce lipid content – e.g. reduction of oil in oats would improve shelf life.

In yeast, especially, oleaginous yeasts, oil content has been increased by choosing a specific carbon and nitrogen source (Ratledge, [15] 1988). The lipid composition has also been changed, for example, in the production of cocoa butter by providing the desired fatty acid in the growth medium as a carbon source (Gierhart, [16] 1984). However, in our process, lipid content has been increased just by enhancing the activity of the first enzyme involved in TAG biosynthesis. It appears that more oil is being produced by creation of a sink by higher GPAT activity. Therefore, no manipulation of growth medium, growth conditions or substrates is required to achieve higher lipid content using this process. We also showed that the GPAT has been effective in increasing total lipids when two types of acyl-CoA were used. It is likely that these GPATs will also be able to use other acyl-CoAs to increase oil content. This will be especially useful in enhancing GLA (Sinden, [17] 1987) or other specialty lipids in yeast. Alternatively, overexpression of GPATs having a selectivity for specific fatty acids could be used to enhance the content of specialty oils in living organisms. The use of lipid enhancing genes such as GPAT may also facilitate the production of commodity oils in microorganisms in an economically sustainable way, a bottleneck thus far (Ratledge, [15] 1988).

SUMMARY

An increase in triglyceride synthesis following expression in yeast of an *E. coli* GPAT and a modified safflower GPAT has been demonstrated by (a) and *in*

vitro assay, (b) NMR analysis of intact yeast and (c) by GC analysis of extracted lipids.

Table 1. Formation of lyso-phosphatidic acid by GPATs

Construct	With 18:1-CoA (pmol/min/mg)		With 16:0-CoA (pmol/min/mg)	
	Glucose	Galactose	Glucose	Galactose
pYES2 (vector)	10.0	6.8	192	131
pYES2:ctpgpat—tp	9.6	33.0	229	175
pYES2:ctpgpat—tp+errs	13.7	44.4	138	157
pYES2: plsB	10.2	28.5	196	790
pYES2: plsB+errs	12.4	20.1	211	346

Table 2. Formation of total lipids by GPATs *in vitro*

Construct	With 18:1-CoA (pmol/min/mg)			With 16:0-CoA (pmol/min/mg)		
	Glucose	Galactose	Ratio*	Glucose	Galactose	Ratio*
pYES2 (vector)	116	112	0.96	1140	513	0.44
pYES2: ctpgpat—tp	184	395	2.15	1170	868	0.74
pYES2: ctpgpat—tp + errs	175	341	1.95	717	771	1.07
pYES2: plsB	98	247	2.51	789	2990	3.77
pYES2: plsB+errs	152	182	1.19	1250	1830	1.46

*galactose/glucose

Table 3. Estimation of oil in transformed yeasts by NMR

Construct	^{1}H-NMR Response*	Ratio of oil contents	% Increase
INVSc (yeast strain)	0.81		
pYES2 (vector)	1.12	1.00	-
pYES2: ctpgpat-tp	1.3	1.17	17
pYES2:ctpgpat-tp+errs	1.68	1.51	51
pYES2: plsB	1.52	1.37	37
pYES2: plsB+errs	1.22	1.10	10

*Cells grown in galactose/Cells grown in glucose
NB: values represent ratios relative to pYES2

Table 4. GC Determination of oil content of transformed yeasts (Wt %)

INVSC1		pYES2		ctpgpat-tp		ctpgpat-tp + errs		plsB		plsB+errs	
a	b	a	b	a	b	a	b	a	b	a	b
4.4	4.4	5.0	6.2	5.5	8.8	6.0	7.4	7.5	7.6	5.0	6.2

Data represent the average of two experiments

a and b represent, respectively, transformed cells grown on a glucose medium and cells in which the promoter is induced using a galactose medium. INVSC1 and pYES 2 represent, respectively, wild type yeast and wild type transformed only with the pYES2 plasmid. These constitute two types of control.

NB: Comparisons can be made only between the a and b sets of data because it cannot be guaranteed that the cells are at an identical growth stage in all experiments.

As far as we are aware there are no references in the published literature dealing directly with the modification of oil content and composition by manipulating the expression of a GPAT. However, selected references dealing with increasing seed oil content by the manipulation of other genes are as follows:

1. Zou J., Katavic V., Giblin EM, Barton DL, Mackenzie SL, Keller WA, Hu X., Taylor DC. 1997. Modification of seed oil content and acyl composition in the Brassicaceae by expression of a yeast *sn*-2 acyltransferase gene. Plant Cell 9:909-923.
2. Zou, J-T, Marillia, E-F, Qi, Q, Barton, DL and Taylor, DC. 1998. Does mitochondrially-generated acetate contribute to plastidial fatty acid biosynthesis? – Antisense repression of an *Arabidopsis thaliana* mitochondrial pyruvate dehydrogenase kinase (PDHK) gene and its effects on oil content and plant development. B-71. 13th International Symposium on Plant Lipids. Seville, Spain, July, 1998.

Other References

3. Frentzen M. 1993. Acyltransferases and Triacylglycerols. In Lipid Metabolism in Plants, Ed TS Moore, Jr. CRC Press, Inc. FL, pp 1195-230.
4. Frentzen M and Wolter FP. 1998. Molecular Biology of Acyltransferanses Involved in glycerolipid Synthesis. In Plant Lipid Biosyntheses: fundamentals and agricultural applications (Ed. JL Harwood). Cambridge University press. pp. 247-272.
5. Wilkison WO and Bell RM 1997. *sn*-Glycerol-3-phosphate acyltransferase from *Escherichia coli*. Biochim Biophys Acta 1348: 3-9.
6. Dircks LK and Sul HS. 1997. Mammalian mitochondrial glycerol-3-phosphate acyltransferase. Biophys Acta 1348: 17-26.

7. Murata N, Ishizaki-Nishizawa O, Higashi S, Hayashi H, Tasaka Y and Nishida I. 1992. Genetically engineered alteration in the chilling sensitivity of plants. Nature 356: 710-713.

8. Nishizawa O. 1996. Chilling-resistant plants and their production. US patent 5,516,667.

9. Wolter FP, Schmidt R and Heinz E. 1992. Chilling sensitivity of Arabidopsis thaliana with genetically engineered membrane lipids. EMBO J 11: 4685-92.

10. Bhella RS and Mackenzie SL. 1994. Nucleotide sequence of a cDNA from *Carthamus tinctorius* encoding a glycerol-3-phosphate acyl transferase. Plant Physiol 106: 1713-1714.

11. Jackson MR, Nilsson T, and Peterson PA. 1990. Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. EMBO J 9: 3153-3162.

12. Eccleston VS and Harwood JL. 1995. Solubilisation, partial purification and properties of acyl-CoA:glycerol-3-phosphate acyltransferase from avocado (Persea americana) fruit mesocarp. Biochimica Biophysica Acta 1257: 1-10.

13. Rutar V. 1989. Magic angle sample spinning NMR spectroscopy of liquids as a nondestructive method for studies of plant seeds. J. Agric Food Chem 37: 70-74.

14. Jamadar SC, Cao WF and Samaniego E. 1996. Relationship between adipose polyamine concentrations and triacylglycerol synthetic enzymes in lean and obese Zucker rats. Enzyme Protein:49:222-230.

15. Ratledge C. 1988. Yeasts for lipid production. Biochem Soc Trans 16: 1088-1091.

16. Gierhart DL. 1984. Multistage process for the preparation of fats and oils. United States patent 4,485,172.

17. Sinden KW. 1987. The production of lipids by fermentation within the EEC. Enzyme Microb Technol 9: 124-125.

What is Claimed is:

1. A method of increasing the oil content and/or the triglyceride composition in living organisms, comprising over-expressing a gene for the enzyme glycerol-3-phosphate acyltransferase, (GPAT).

2. A method according to Claim 1, wherein the living organism is yeast.

3. A method according to Claim 1, wherein the GPAT gene has the nucleotide sequence of Figure 1.

4. A method according to Claim 1, wherein the gene has the nucleotide sequence of Figure 3.

5. A method according to Claim 1, wherein the gene has the nucleotide sequence of Figure 5.

6. A method according to Claim 1, wherein the gene has the nucleotide sequence of Figure 7.

GGATCCATGCACGGTCACTCTCGTACATTCATCGATGCTCGTTCCGAACAAGATCTTCTTTCTGGAATTCAAAGAGAGTT
GGAAGCTGGAACACTGCCAAAACATATTGCTCAAGCAATGGAGGAGCTATATCAGAACTACAAAAATGCAGTTCTCCAAA
GTGCGGCTCCTCATGCAGAAGATATTGTGTTGTCAAACATGCGTGTAGCGTTTGATCGTATGTTCTTGGATGTGAAGGAG
CCGTTTGAATTTTCACCATATCATGAAGCTATTTTGGAACCTTTTAACTACTATATGTTTGGTCAAAATTATATTCGGCC
TTTGGTCAATTTCAGGGAATCATACGTTGGCAATGTCTCCGTTTTCGGTGTAATGGAAGAGCAGCTTAAGCAGGGTGACA
AGGTGGTTTTGATCTCAAACCATCAAACAGAAGCAGATCCAGCTGTTATTGCCTTGATGCTTGAAACAACAAACCCCCAT
ATTTCTGAGAACATAATCTACGTGGCAGGGGATAGAGTAATAACAGATCCTCTTTGCAAGCCTTTCAGCATGGGAAGGAA
TCTGTTGTGCGTGTATTCAAAAAAGCATATGAATGATGTTCCTGAGCTTGCTGAGATGAAAAAAAGATCAAATACAAGAA
GTTTAAAAGGGAGGATGGCTTTGCTTTTGAGGGGCGGATCTAAAATAATATGGATTGCGCCAAGTGGTGGCAGGGACAGG
CCAGATCCTATCACAAATCAGTGGTTTCCGGCACCGTTTGATGCCACTTCGCTTGACAACATGAGAAGGCTCGTGGACCA
TGCTGGTTTGGTGGGTCACATATATCCTTTAGCCATATTGTGCCATGACATCATGCCCCCTCCTCTTCAGGTTGAGAAAG
AAATTGGAGAGAAGAGTTGGATCTCTTTTCATGGCACCGGAATATCAGTGGCACCGGAAATTAATTTCCAAGAAGTTACT
GCCTCTTGTGGGTCCCCCGAGGAGGCGAAGGCAGCTTATTCACAGGCACTCTATGATTCCGTGTGTGAACAATACAAGGT
GCTACATTCTGCGGTACATGGAGGAAAAGGGTTAGAAGCATCAACACCAAGTGTCTCGTTGTCACAACCCTTGCAGTTTC
TCGATTAGGATCC

Figure 1

Nucleotide sequence of the gene *ctpgpat–tp* (*Carthamus tinctorius* plastidial GPAT from which the sequence for a putative transit peptide was removed) used for transforming yeast. Start and stop codon is underlined. ATG was introduced as a start codon in place of GAC.

GSMHGHSRTFIDARSEQDLLSGIQRELEAGTLPKHIAQAMEELYQNYKNAVLQSAAPHAEDIVLSNMRVAFDRMFLDVKE
PFEFSPYHEAILEPFNYYMFGQNYIRPLVNFRESYVGNVSVPGVMEEQLKQGDKVVLISNHQTEADPAVIALMLETTNPH
ISENIIYVAGDRVITDPLCKPFSMGRNLLCVYSKKHMNDVPELAEMKKRSNTRSLKGRMALLLRGGSKIIWIAPSGGRDR
PDPITNQWFPAPFDATSLDNMRRLVDHAGLVGHIYPLAILCHDIMPPPLQVEKEIGEKSWISFHGTGISVAPEINFQEVT
ASCGSPEEAKAAYSQALYDSVCEQYKVLHSAVHGGKGLEASTPSVSLSQPLQFLD.D

Figure 2

Translated protein sequence of the gene *ctpgpat–tp* used for transforming yeast. The mature polypeptide will not have the first 90 amino acids of the ctgpat (Bhella and MacKenzie, [10] 1994). Amino acid 91, Asp (D), is replaced with Met (M).

GGATCCATGCACGGTCACTCTCGTACATTCATCGATGCTCGTTCCGAACAAGATCTTCTTTCTGGAATTCAAAGAGAGTT
GGAAGCTGGAACACTGCCAAAACATATTGCTCAAGCAATGGAGGAGCTATATCAGAACTACAAAAATGCAGTTCTCCAAA
GTGCGGCTCCTCATGCAGAAGATATTGTGTTGTCAAACATGCGTGTAGCGTTTGATCGTATGTTCTTGGATGTGAAGGAG
CCGTTTGAATTTTCACCATATCATGAAGCTATTTTGGAACCTTTTAACTACTATATGTTTGGTCAAAATTATATTCGGCC
TTTGGTCAATTTCAGGGAATCATACGTTGGCAATGTCTCCGTTTTCGGTGTAATGGAAGAGCAGCTTAAGCAGGGTGACA
AGGTGGTTTTGATCTCAAACCATCAAACAGAAGCAGATCCAGCTGTTATTGCCTTGATGCTTGAAACAACAAACCCCCAT
ATTTCTGAGAACATAATCTACGTGGCAGGGGATAGAGTAATAACAGATCCTCTTTGCAAGCCTTTCAGCATGGGAAGGAA
TCTGTTGTGCCGTGTATTCAAAAAAGCATATGAATGATGTTCCTGAGCTTGCTGAGATGAAAAAAAGATCAAATACAAGAA
GTTTAAAAGGGAGGATGGCTTTGCTTTTGAGGGGCGGATCTAAAATAATATGGATTGCGCCAAGTGGTGGCAGGGACAGG
CCAGATCCTATCACAAATCAGTGGTTTCCGGCACCGTTTGATGCCACTTCGCTTGACAACATGAGAAGGCTCGTGGACCA
TGCTGGTTTGGTGGGTCACATATATCCTTTAGCCATATTGTGCCATGACATCATGCCCCCTCCTCTTCAGGTTGAGAAAG
AAATTGGAGAGAAGAGTTGGATCTCTTTTCATGGCACCGGAATATCAGTGGCACCGGAAATTAATTTCCAAGAAGTTACT
GCCTCTTGTGGGTCCCCCGAGGAGGCGAAGGCAGCTTATTCACAGGCACTCTATGATTCCGTGTGTGAACAATACAAGGT
GCTACATTCTGCGGTACATGGAGGAAAAGGGTTAGAAGCATCAACACCAAGTGTCTCGTTGTCACAACCCAAGCAGAAAC
TCGATTAGGATCC

Figure 3

Nucleotide sequence of the gene *ctpgpat–tp+errs* (*Carthamus tinctorius* plastidial GPAT from which the sequence for a putative transit peptide was removed and sequences for endoplasmic reticulum retention signals were added). Start and stop codons are underlined. ATG was introduced as a start codon in place of GAC. A dditionally, AAG and AAA (also underlined) in place of TTG and TTT were introduced for ER retention of protein.

GSMHGHSRTFIDARSEQDLLSGIQRELEAGTLPKHIAQAMEELYQNYKNAVLQSAAPHAEDIVLSNMRVAFDRMFLDVKE
PFEFSPYHEAILEPFNYYMFGQNYIRPLVNFRESYVGNVSVFGVMEEQLKQGDKVVLISNHQTEADPAVIALMLETTNPH
ISENIIYVAGDRVITDPLCKPFSMGRNLLCVYSKKHMNDVPELAEMKKRSNTRSLKGRMALLLRGGSKIIWIAPSGGRDR
PDPITNQWFPAPFDATSLDNMRRLVDHAGLVGHIYPLAILCHDIMPPPLQVEKEIGEKSWISFHGTGISVAPEINFQEVT
ASCGSPEEAKAAYSQALYDSVCEQYKVLHSAVHGGKGLEASTPSVSLSQPKQKLD.D

Figure 4

Translated protein sequence of the gene *ctpgpat–tp+errs*. The mature polypeptide will not have the first 90 amino acids of the *ctgpat* (Bhella and MacKenzie, 1994). Amino acid 91, Asp (D), is replaced with Met (M). Also Lys (K), underlined, replaced a Phe (F) and Leu (L) at position 3 and 5 from the C terminus.

AGATCTTCCCATGACTTTCTGCTATCCTTGCCGCGCATTTGCATTATTAACCAGAGGCTTTACATCGTTTATGTCCGGCT
GGCCACGAATTTACTACAAATTACTGAATTTACCATTAAGCATCCTGGTAAAAAGCAAGTCTATTCCGGCAGATCCTGCC
CCGGAACTGGGGCTGGATACCTCTCGTCCAATTATGTACGTTTTACCGTACAACTCGAAAGCAGATTTGCTGACGTTGCG
CGCCCAGTGTCTGGCACATGACTTGCCTGACCCGTTAGAGCCGCTGGAAATCGACGGCACGCTACTGCCGCGCTATGTGT
TCATTCACGGCGGGCCGCGTGTGTTCACCTATTACACGCCGAAAGAAGAGTCTATTAAGCTGTTCCACGACTATCTCGAT
TTGCACCGTAGCAACCCAAATCTGGATGTGCAGATGGTGCCAGTGTCGGTGATGTTTGGTCGCGCGCCGGGGCGTGAAAA
AGGCGAAGTGAACCCGCCGCTGCGTATGCTTAACGGCGTACAGAAATTTTTCGCTGTACTGTGGCTCGGTCGCGACAGTT
TTGTGCGTTTCTCGCCGTCAGTTTCGCTGCGCCGTATGGCGGATGAACACGGCACGGATAAAACTATCGCTCAGAAACTG
GCGCGCGTGGCGCGTATGCACTTTGCCCGTCAACGTCTGGCTGCCGTAGGCCCACGTTTGCCTGCTCGTCAGGATCTGTT
TAATAAGCTGCTCGCCTCCCGCGCCATTGCCAAAGCGGTAGAAGATGAAGCGCGCAGCAAAAAAATCTCCCATGAAAAAG
CGCAGCAGAACGCGATTGCACTGATGGAAGAGATTGCGGCGAATTTCTCTTACGAGATGATTCGCCTGACTGACCGTATT
CTGGGCTTCACCTGGAACCGACTTTACCAGGGCATCAACGTCCATAACGCTGAGCGCGTTCGCCAGCTGGCCCACGACGG
TCATGAGCTGGTATATGTGCCTTGCCACCGCAGTCACATGGACTACCTGCTGCTTTCTTACGTGCTGTATCACCAGGGGC
TGGTGCCGCCGCATATCGCCGCCGGGATCAACCTGAATTTCTGGCCTGCCGGGCCGATTTTCCGCCGTCTGGGGGCGTTC
TTTATTCGCCGTACGTTTAAAGGCAATAAACTTTATTCCACCGTTTTCCGGGAGTATCTCGGCGAACTGTTCAGCCGTGG
TTATTCCGTCGAGTACTTCGTGGAAGGCGGTCGTTCCCGTACGGGGCGTTTGCTGGATCCGAAAACTGGTACGCTGTCGA
TGACCATTCAGGCGATGCTGCGTGGCGGCACGCGTCCGATTACGCTGATTCCGATCTATATCGGTTATGAGCACGTCATG
GAAGTGGGTACTTACGCCAAAGAACTGCGCGGCGCGACGAAAGAGAAAGAGAGCCTGCCGCAGATGCTGCGCGGTTTAAG
CAAGCTGCGTAATCTCGGTCAGGGTTACGTCAACTTCGGTGAACCAATGCCGTTGATGACCTACCTTAACCAGCATGTAC
CTGACTGGCGTGAATCTATCGATCCCATCGAAGCGGTGCGTCCGGCATGGTTAACGCCGACGGTCAATAATATTGCTGCC
GATCTGATGGTACGCATTAACAACGCAGGCGCGGCAAACGCCATGAACCTGTGCTGTACTGCGCTACTGGCATCACGTCA
GCGCTCACTCACCCGCGAGCAGTTAACCGAGCAACTCAACTGCTACCTGGATCTGATGCGCAACGTGCCCTACTCCACGG
ACTCTACCGTTCCTTCAGCCAGCGCCAGCGAGCTTATCGATCACGCGCTGCAAATGAACAAGTTTGAAGTCGAGAAAGAC
ACAATCGGCGACATCATCATTCTGCCGCGCGAGCAAGCGGTGCTGATGACCTACTATCGCAACAACATTGCGCATATGTT
GGTGCTGCCTTCGCTGATGGCGGCAATCGTCACCCAGCATCGCCACATCTCCCGCGACGTATTGATGGAGCACGTCAATG
TGCTTTACCCAATGCTGAAAGCGGAGCTGTTCCTGCGCTGGGATCGCGACGAGTTGCCGGACGTTATTGATGCGCTGGCA
AATGAGATGCAACGTCAGGGGCTGATTACCCTGCAAGATGATGAGTTGCATATCAACCCGGCGCATTCTCGCACGCTACA
GCTGCTGGCCGCAGGCGCGCGCGAAACGCTGCAACGTTATGCCATCACCTTCTGGTTGTTGAGTGCCAACCCGTCGATCA
ACCGCGGTACGCTGGAGAAAGAGAGCCGCACCGTCGCGCAACGTCTCTCCGTGCTGCACGGCATCAACGCGCCGGAGTTC
TTCGACAAGGCGGTGTTCAGTTCTCTGGTGCTGACACTGCGTGATGAAGGGTATATCAGCGATAGCGGCGATGCCGAACC
GGCAGAAACGATGAAGGTTTATCAGTTGCTGGCGGAGTTGATTACATCAGACGTGCGTTTGACGATTGAGAGTGCGACGC
AGGGCGAAGGGTAATCAGATCT

Figure 5

Nucleotide sequence of the gene *plsB* (*E.coli* GPAT). Start and stop codon is underlined.

MTFCYPCRAFALLTRGFTSFMSGWPRIYYKLLNLPLSILVKSKSIPADPAPELGLDTSRPIMYVLPYNSKADLLTLRAQC
LAHDLPDPLEPLEIDGTLLPRYVFIHGGPRVFTYYTPKEESIKLFHDYLDLHRSNPNLDVQMVPVSVMFGRAPGREKGEV
NPPLRMLNGVQKFFAVLWLGRDSFVRFSPSVSLRRMADEHGTDKTIAQKLARVARMHFARQRLAAVGPRLPARQDLFNKL
LASRAIAKAVEDEARSKKISHEKAQQNAIALMEEIAANFSYEMIRLTDRILGFTWNRLYQGINVHNAERVRQLAHDGHEL
VYVPCHRSHMDYLLLSYVLYHQGLVPPHIAAGINLNFWPAGPIFRRLGAFFIRRTFKGNKLYSTVFREYLGELFSRGYSV
EYFVEGGRSRTGRLLDPKTGTLSMTIQAMLRGGTRPITLIPIYIGYEHVMEVGTYAKELRGATKEKESLPQMLRGLSKLR
NLGQGYVNFGEPMPLMTYLNQHVPDWRESIDPIEAVRPAWLTPTVNNIAADLMVRINNAGAANAMNLCCTALLASRQRSL
TREQLTEQLNCYLDLMRNVPYSTDSTVPSASASELIDHALQMNKFEVEKDTIGDIIILPREQAVLMTYYRNNIAHMLVLP
SLMAAIVTQHRHISRDVLMEHVNVLYPMLKAELFLRWDRDELPDVIDALANEMQRQGLITLQDDELHINPAHSRTLQLLA
AGARETLQRYAITFWLLSANPSINRGTLEKESRTVAQRLSVLHGINAPEFFDKAVFSSLVLTLRDEGYISDSGDAEPAET
MKVYQLLAELITSDVRLTIESATQGEG.SD

Figure 6

Translated protein sequence of the gene *plsB* (*E.coli* GPAT).

AGATCTTCCCATGACTTTCTGCTATCCTTGCCGCGCATTTGCATTATTAACCAGAGGCTTTACATCGTTTATGTCCGGCT
GGCCACGAATTTACTACAAATTACTGAATTTACCATTAAGCATCCTGGTAAAAAGCAAGTCTATTCCGGCAGATCCTGCC
CCGGAACTGGGGCTGGATACCTCTCGTCCAATTATGTACGTTTTACCGTACAACTCGAAAGCAGATTTGCTGACGTTGCG
CGCCCAGTGTCTGGCACATGACTTGCCTGACCCGTTAGAGCCGCTGGAAATCGACGGCACGCTACTGCCGCGCTATGTGT
TCATTCACGGCGGGCCGCGTGTGTTCACCTATTACACGCCGAAAGAAGAGTCTATTAAGCTGTTCCACGACTATCTCGAT
TTGCACCGTAGCAACCCAAATCTGGATGTGCAGATGGTGCCAGTGTCGGTGATGTTTGGTCGCGCGCCGGGGCGTGAAAA
AGGCGAAGTGAACCCGCCGCTGCGTATGCTTAACGGCGTACAGAAATTTTTCGCTGTACTGTGGCTCGGTCGCGACAGTT
TTGTGCGTTTCTCGCCGTCAGTTTCGCTGCGCCGTATGGCGGATGAACACGGCACGGATAAAACTATCGCTCAGAAACTG
GCGCGCGTGGCGCGTATGCACTTTGCCCGTCAACGTCTGGCTGCCGTAGGCCCACGTTTGCCTGCTCGTCAGGATCTGTT
TAATAAGCTGCTCGCCTCCCGCGCCATTGCCAAAGCGGTAGAAGATGAAGCGCGCAGCAAAAAAATCTCCCATGAAAAAG
CGCAGCAGAACGCGATTGCACTGATGGAAGAGATTGCGGCGAATTTCTCTTACGAGATGATTCGCCTGACTGACCGTATT
CTGGGCTTCACCTGGAACCGACTTTACCAGGGCATCAACGTCCATAACGCTGAGCGCGTTCGCCAGCTGGCCCACGACGG
TCATGAGCTGGTATATGTGCCTTGCCACCGCAGTCACATGGACTACCTGCTGCTTTCTTACGTGCTGTATCACCAGGGGC
TGGTGCCGCCGCATATCGCCGCCGGGATCAACCTGAATTTCTGGCCTGCCGGGCCGATTTTCCGCCGTCTGGGGGCGTTC
TTTATTCGCCGTACGTTTAAAGGCAATAAACTTTATTCCACCGTTTTCCGGGAGTATCTCGGCGAACTGTTCAGCCGTGG
TTATTCCGTCGAGTACTTCGTGGAAGGCGGTCGTTCCCGTACGGGGCGTTTGCTGGATCCGAAAACTGGTACGCTGTCGA
TGACCATTCAGGCGATGCTGCGTGGCGGCACGCGTCCGATTACGCTGATTCCGATCTATATCGGTTATGAGCACGTCATG
GAAGTGGGTACTTACGCCAAAGAACTGCGCGGCGCGACGAAAGAGAAAGAGAGCCTGCCGCAGATGCTGCGCGGTTTAAG
CAAGCTGCGTAATCTCGGTCAGGGTTACGTCAACTTCGGTGAACCAATGCCGTTGATGACCTACCTTAACCAGCATGTAC
CTGACTGGCGTGAATCTATCGATCCCATCGAAGCGGTGCGTCCGGCATGGTTAACGCCGACGGTCAATAATATTGCTGCC
GATCTGATGGTACGCATTAACAACGCAGGCGCGGCAAACGCCATGAACCTGTGCTGTACTGCGCTACTGGCATCACGTCA
GCGCTCACTCACCCGCGAGCAGTTAACCGAGCAACTCAACTGCTACCTGGATCTGATGCGCAACGTGCCCTACTCCACGG
ACTCTACCGTTCCTTCAGCCAGCGCCAGCGAGCTTATCGATCACGCGCTGCAAATGAACAAGTTTGAAGTCGAGAAAGAC
ACAATCGGCGACATCATCATTCTGCCGCGCGAGCAAGCGGTGCTGATGACCTACTATCGCAACAACATTGCGCATATGTT
GGTGCTGCCTTCGCTGATGGCGGCAATCGTCACCCAGCATCGCCACATCTCCCGCGACGTATTGATGGAGCACGTCAATG
TGCTTTACCCAATGCTGAAAGCGGAGCTGTTCCTGCGCTGGGATCGCGACGAGTTGCCGGACGTTATTGATGCGCTGGCA
AATGAGATGCAACGTCAGGGGCTGATTACCCTGCAAGATGATGAGTTGCATATCAACCCGGCGCATTCTCGCACGCTACA
GCTGCTGGCCGCAGGCGCGCGCGAAACGCTGCAACGTTATGCCATCACCTTCTGGTTGTTGAGTGCCAACCCGTCGATCA
ACCGCGGTACGCTGGAGAAAGAGAGCCGCACCGTCGCGCAACGTCTCTCCGTGCTGCACGGCATCAACGCGCCGGAGTTC
TTCGACAAGGCGGTGTTCAGTTCTCTGGTGCTGACACTGCGTGATGAAGGGTATATCAGCGATAGCGGCGATGCCGAACC
GGCAGAAACGATGAAGGTTTATCAGTTGCTGGCGGAGTTGATTACATCAGACGTGCGTTTGACGATTGAGAGTGCGAAGC
AGAAGGAAGGGTAATCAGATCT

Figure 7

Nucleotide sequence of the gene *plsB+errs* (*E.coli* GPAT to which sequences for endoplasmic reticulum retention signals were added). Start and stop codons are underlined. AAG and AAG (also underlined) in place of GGC and ACG were introduced at the 3' end for ER retention of protein.

MTFCYPCRAFALLTRGFTSFMSGWPRIYYKLLNLPLSILVKSKSIPADPAPELGLDTSRPIMYVLPYNSKADLLTLRAQC
LAHDLPDPLEPLEIDGTLLPRYVFIHGGPRVFTYYTPKEESIKLFHDYLDLHRSNPNLDVQMVPVSVMFGRAPGREKGEV
NPPLRMLNGVQKFFAVLWLGRDSFVRFSPSVSLRRMADEHGTDKTIAQKLARVARMHFARQRLAAVGPRLPARQDLFNKL
LASRAIAKAVEDEARSKKISHEKAQQNAIALMEEIAANFSYEMIRLTDRILGFTWNRLYQGINVHNAERVRQLAHDGHEL
VYVPCHRSHMDYLLLSYVLYHQGLVPPHIAAGINLNFWPAGPIFRRLGAFFIRRTFKGNKLYSTVFREYLGELFSRGYSV
EYFVEGGRSRTGRLLDPKTGTLSMTIQAMLRGGTRPITLIPIYIGYEHVMEVGTYAKELRGATKEKESLPQMLRGLSKLR
NLGQGYVNFGEPMPLMTYLNQHVPDWRESIDPIEAVRPAWLTPTVNNIAADLMVRINNAGAANAMNLCCTALLASRQRSL
TREQLTEQLNCYLDLMRNVPYSTDSTVPSASASELIDHALQMNKFEVEKDTIGDIIILPREQAVLMTYYRNNIAHMLVLP
SLMAAIVTQHRHISRDVLMEHVNVLYPMLKAELFLRWDRDELPDVIDALANEMQRQGLITLQDDELHINPAHSRTLQLLA
AGARETLQRYAITFWLLSANPSINRGTLEKESRTVAQRLSVLHGINAPEFFDKAVFSSLVLTLRDEGYISDSGDAEPAET
MKVYQLLAELITSDVRLTIESAKQKEG.SD

Figure 8

Translated protein sequence of the gene *plsB+errs*. Amino acid Lys (K), underlined, replaced a Gly (G) and Thr (T) at position 3 and 5 from the C terminus.